SHOCK, Vol. 42, No. 2, pp. 108–114, 2014

UTILITY OF VITAL SIGNS, HEART RATE VARIABILITY AND COMPLEXITY, AND MACHINE LEARNING FOR IDENTIFYING THE NEED FOR LIFESAVING INTERVENTIONS IN TRAUMA PATIENTS

Nehemiah T. Liu,* John B. Holcomb,† Charles E. Wade,† Mark I. Darrah,‡ and Jose Salinas*

**US Army Institute of Surgical Research, Fort Sam Houston; and †Center for Translational Injury Research, Department of Surgery, University of Texas Health Science Center at Houston, Houston, Texas; and ‡Athena GTX, Inc, Des Moines, Iowa*

Received 20 Feb 2014; first review completed 17 Mar 2014; accepted in final form 31 Mar 2014

ABSTRACT—To date, no studies have attempted to utilize data from a combination of vital signs, heart rate variability and complexity (HRV, HRC), as well as machine learning (ML), for identifying the need for lifesaving interventions (LSIs) in trauma patients. The objectives of this study were to examine the utility of the above for identifying LSI needs and compare different LSI-associated models, with the hypothesis that an ML model would be superior in performance over multivariate logistic regression models. One hundred four patients transported from the injury scene via helicopter were selected for the study. A wireless vital signs monitor was attached to the patient's arm and used to capture physiologic data, including HRV and HRC. The power of vital sign measurements, HRV, HRC, and Glasgow Coma Scale score (GCS) to identify patients requiring LSIs was estimated using multivariate logistic regression and ML. Receiver operating characteristic (ROC) curves were also obtained. Thirty-two patients underwent 75 LSIs. After logistic regression, ROC curves demonstrated better identification for LSIs using heart rate (HR) and HRC (area under the curve [AUC] of 0.81) than using HR alone (AUC of 0.73). Likewise, ROC curves demonstrated better identification for LSIs using GCS and HRC (AUC of 0.94) than using GCS and HR (AUC of 0.92). Importantly, ROC curves demonstrated that an ML model using HR, GCS, and HRC (AUC of 0.99) had superior performance over multivariate logistic regression models for identifying the need for LSIs in trauma patients. Development of computer decision support systems should utilize vital signs, HRC, and ML in order to achieve more accurate diagnostic capabilities, such as identification of needs for LSIs in trauma patients.

KEYWORDS—Machine learning, lifesaving interventions, heart rate complexity, heart rate variability, trauma

ABBREVIATIONS—AUC — area under the curve; CDS — computer decision support; CI — confidence interval; DBP — diastolic blood pressure; ECG — electrocardiogram; ED — emergency department; GCS — Glasgow Coma Scale; HF — high frequency; HR — heart rate; HRC — heart rate complexity; HRV — heart rate variability; LF — low frequency; LSI — lifesaving intervention; MAP — mean arterial pressure; ML — machine learning; ROC — receiver operating characteristic; RR — respiratory rate; SampEn — sample entropy; SI — Shock index; SPo_2 — blood oxygenation; WVSM — Wireless Vital Signs Monitor

INTRODUCTION

Capture of high-frequency (HF) data for real-time triage and assessment of trauma patients is now a viable option due to advances in sensor technology and computing power in mobile platforms (1, 2). These advances will allow for a new generation of information-driven computer decision support (CDS) systems that could significantly enhance medical decision making and lead to improvements in outcome (2–5). Still, in order to achieve more accurate diagnostic capabilities, development of these systems may require new approaches based on combinations of standard vital signs, trends, and signal-derived metrics, fused with advanced artificial intelligence or machine learning (ML) technologies (2, 6, 7).

Previous studies have shown that standard vital signs alone may not be reliable for timely and accurate assessment of true injury severity in trauma patients because of erroneous measurements or inherent physiologic compensatory mechanisms, which could lead to errors in diagnosis (8–12). However, use of new advanced indices derived from the electrocardiogram (ECG)—namely, heart rate variability and complexity (HRV, HRC)—may provide one alternative for monitoring trauma patients more reliably and accurately (13–18). HRV and HRC metrics have been shown to be useful not only for detecting acute changes in patient stability (13, 18) but also for risk stratification (15, 16) and identification of patients requiring lifesaving interventions (LSIs) (14, 17). Furthermore, they are noninvasive and can be calculated via automation and telemetry within seconds (16, 17, 19). Whereas vital signs may originate from a single source of failure, streaming HRV and HRC values can be obtained from multiple asynchronous waveform sources (20). Nevertheless, HRV and HRC have several significant drawbacks, including limited use in the

Address reprint requests to Nehemiah T. Liu, MS, US Army Institute of Surgical Research, 3698 Chambers Pass, JBSA Fort Sam Houston, TX 78234-6315. E-mail: nehemiah.liu@us.army.mil.

Author Contributions: N.T.L. contributed to study design, data analysis, data interpretation, writing, and critical revision. J.B.H., C.E.W., and M.I.D. contributed to study design and critical revision. J.S. contributed to study design, writing, and critical revision.

This work was supported by the National Trauma Institute, the Combat Casualty Care Research Program, and the State of Texas Emerging Technology Fund.

The authors declare no conflicts of interest.

This study was conducted under a protocol reviewed and approved by the University of Texas Health Science Center at Houston and in accordance with the approved protocol. The opinions or assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the Department of the Army or the Department of Defense.

DOI: 10.1097/SHK.0000000000000186

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **01 AUG 2014**	2. REPORT TYPE **N/A**	3. DATES COVERED **-**

4. TITLE AND SUBTITLE **Utility of vital signs, heart-rate variability and complexity, and machine learning for identifying the need for life-saving interventions in trauma patients**	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **US Army Institute of Surgical Research, JBSA Fort Sam Houston, Texas**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**UU**	**7**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

presence of high noise levels in the captured waveforms (20–23). Therefore, development of advanced CDS systems may need to utilize combinations of standard vital signs in addition to HRV and HRC indices in order to provide better diagnosis capabilities in complex patient care environments and mitigate issues related to processing of individual data sources.

In addition, because medical decision making is highly complex, subjective, and difficult to predict (6, 17, 24), new CDS systems need to incorporate more objective tools such as artificial intelligence and ML into the decision-making process to effectively process and fuse multiple heterogeneous data sources generated by the patient. By modeling relationships between seemingly disparate parameters and outcomes using ML techniques, CDS systems could not only equip providers in the intensive care unit with more accurate, actionable information, such as which potential treatments to prescribe (4, 5) or when to perform LSIs (6), but also enhance the medic's ability to assess battlefield injuries (automated triage) (2, 6).

Although the last few decades have witnessed an emergence of various CDS systems and studies investigating their clinical relevance/performance (4, 5), to date, no studies have attempted to utilize data from a combination of standard vital signs, HRV, and HRC, as well as ML, for purposes of identifying needs for LSIs in trauma patients, nor have studies attempted to compare different LSI identification models for different injury patterns. The aims of this study were to (*a*) confirm that HRV and HRC can discriminate between those patients who received one or more LSIs and those who received none and (*b*) examine the efficacy of a combined model of standard vital signs, HRV, and HRC, for predicting the need for LSIs in trauma patients using both multivariate regression modeling and ML-based modeling. Our hypothesis was that an ML system utilizing vital signs, HRV, and HRC to identify the needs for LSIs would be able to outperform standard statistically derived multivariate logistic regression models.

MATERIALS AND METHODS

Subjects and protocol

This study was approved by the institutional review board at the University of Texas Health Science Center, Houston, Tex. The data set used in this study consisted of a convenience cohort of 104 patients transported via the Life Flight helicopter service to the Memorial Hermann Hospital, a level I trauma center in Houston, Tex, between June 27, 2011, and January 6, 2012. All patients were prehospital trauma patients, and all wore a Wireless Vital Signs Monitor (WVSM; Athena GTX, Inc, Des Moines, Iowa) system during transport, admission to the hospital, and stay in the emergency department (ED). The WVSM was used to capture numeric and waveform data, which were then transmitted to a computerized server system via a wireless connection. Thus, both prehospital and ED LSIs were performed during continuous WVSM monitoring.

Numeric data from the WVSM device were stored at a rate of 1 Hz. These data included HRV calculated every second via the method of an HF–low-frequency (LF) power spectrum ratio (25, 26), and HRC calculated every second via the method of sample entropy (SampEn) (27) (see Heart rate variability and complexity). Total Glasgow Coma Scale (GCS) scores were also recorded every second in order for scores to be time synchronized with other numeric data at 1 Hz and updated only upon manual entry from the WVSM device. In other words, all scores were determined manually by physical examination, recorded continuously, and inputted into the device whenever patient status changed.

Single-lead ECG waveform data and plethysmograph waveform data from a thumb-mounted pulse oximeter to the WVSM were recorded at rates of 230 and 75 Hz, respectively. For intubated patients, respiration waveform data were also recorded at a rate of 10 Hz using a handheld capnograph/oximeter (Microcap; Covidien, Mansfield, Mass). Standard vital signs used during trauma care for patient assessment included heart rate (HR), systolic blood pressure (SBP), diastolic blood pressure (DBP), mean arterial pressure (MAP), respiratory rate (RR), and blood oxygenation (SpO_2). Combinations of these vital signs were also used to derive other measurements, including shock index (SI = HR/SBP) and pulse pressure (PP = SBP − DBP).

All nonelectronic data were manually recorded on an electronic run sheet (RescueNet ePCR; Zoll Medical, Chelmsford, Mass) by emergency medical services medics, then collected on a standardized form and entered into a database (OpenClinica, https://www.openclinica.com/). These included demographic data, physical examination results, individual components of GCS scores (motor, eye, verbal), and interventions performed on the patients in the field and ED. Prehospital and ED LSIs consisted of angioembolizations, blood transfusions, cardioversions, cardiopulmonary resuscitations, cricothyrotomies, endotracheal intubations, pericardiocenteses, thoracotomies, tourniquets, tube thoracostomies, and needle decompressions.

Heart rate variability and complexity

For this study, HRV was derived using a fast Fourier transform of the R-R intervals captured from the patient's ECG. An HRV index was computed using the ratio of the HF power (HF: 0.15–0.4 Hz) corresponding to the parasympathetic (vagal) activation to the LF power (LF: 0.04–0.15 Hz) corresponding to the sympathetic and vagal activation (26).

Heart rate complexity quantifies the structural complexity in the R-R interval sequence (i.e., complexity in the patterns of the HR time series) (22, 26–28). For this study, HRC was calculated via the method of SampEn (m, r, N) because of its suitability for analysis of shorter time series. Sample entropy was calculated using the negative natural logarithm of the conditional probability that two epochs similar for m intervals remain similar at the next interval, given a sequence of N intervals and excluding self-matches. In this case, similarity was defined as intervals differing by no more than some tolerance r (in milliseconds) (26). Values of SampEn were obtained by the following equations:

$$SampEn(m,r,N) = -\ln(A/B), \tag{1}$$

$$B = [(N-m-1)/2]\sum_{i=1}^{N-m} B_i^r(m), \tag{2}$$

$$A = [(N-m-1)/2]\sum_{i=1}^{N-m} A_i^r(m). \tag{3}$$

For a sequence of N intervals, if $x_m(i)$ is an epoch of m consecutive intervals starting at index i and running from $i = 1, \ldots, N - m$, then $B^r{}_i(m)$ denotes the number of epochs $x_m(j)$ within r of $x_m(i)$, for $i \neq j$, multiplied by $(N - m - 1)^{-1}$, and $A^r{}_i(m)$ denotes the number of epochs $x_{m+1}(j)$ within r of $x_{m+1}(i)$, for $i \neq j$, multiplied by $(N - m - 1)^{-1}$ (26).

Parametric values (N = 200, m = 2, r = 6) were established from previous work (14, 16, 17). A higher SampEn implies a signal with more complexity as well as a higher likelihood that the signal belongs to a healthy patient (22, 26–28).

Statistical analyses

Normality was not assumed for means within each group and across all records because of the small sample size. All data sets were analyzed using Wilcoxon tests for nonparametric distributions. Initial multivariate logistic regression analyses were also done for all subjects with independent variables of age, height, race, and weight and with dependent variables of HR, SBP, DBP, MAP, RR, and SI. These analyses excluded HRV/HRC values. Factors that were not significant ($P > 0.05$) were removed from the model via backward elimination. A second set of analyses were done for dependent variables of HR, SBP, DBP, MAP, RR, and SI, adding HRV and HRC for performance comparisons with the initial set. In addition, a third and fourth set of analyses were performed for all subjects in order to include GCS scores, with and without HRV and HRC as dependent variables, respectively.

Machine learning analyses and modeling was performed using artificial neural networks and multilayer perceptron models for all subjects with vital signs, HRV, HRC, and GCS scores. Receiver operating characteristic (ROC) curves were obtained to examine the discriminating power of the models for the outcome of at least one LSI.

The accuracy of the statistical models was assessed using sensitivity and specificity scores. The power of demographics, vital sign measurements, HRV, HRC, and GCS scores to identify whether LSIs were performed was estimated using multivariate logistic regression and ML (neural networks, multilayer perceptrons). JMP version 9.0.0 (SAS Institute, Cary, NC) and the R Language (http://www.r-project.org/), a well-known open-source statistical software package, were used for statistical analyses.

RESULTS

Physiologic data were collected on a convenience sample of 104 patients. Patient demographics are shown in Table 1.

Table 1. **Demographics**

	All patients		LSI patients		Prehospital LSIs		ED LSIs	
	#	%	#	%	#	%	#	%
Variable	N	N/104	n	n/N	*i*	*i*/31	*j*	*j*/44
All patients	104	100	32	31	31	100	44	100
Injury type								
Blunt	94	90	29	31	27	87	42	95
Penetrating	10	10	3	30	4	13	2	5
Total GCS*								
Mean 12 ± 5								
3	22	21	21	95	26	84	29	66
4	1	1	1	100	1	4	0	0
13	3	3	0	0	0	0	0	0
14	23	22	3	13	2	6	3	7
15	54	52	6	11	0	0	10	23
Unknown	1	1	1	100	2	6	2	4
Motor GCS*								
Mean 5 ± 2								
1	22	21	21	95	26	84	29	66
2	1	1	1	100	1	4	0	0
6	80	77	9	11	2	6	13	30
Unknown	1	1	1	100	2	6	2	4
Gender								
Female	22	21	6	27	8	26	6	14
Male	82	79	26	32	23	74	38	86
Race								
White/Caucasian	62	60	18	29	17	55	27	62
Black	11	10	3	27	3	10	5	11
Hispanic	23	22	11	50	11	35	12	27
Asian/Pacific	1	1	0	0	0	0	0	0
Not recorded	7	7	0	0	0	0	0	0
Age, y								
Mean 40 ± 16								
Quartiles								
18–26	26	25	6	23	6	19	11	25
27–37	26	25	7	27	7	23	6	14
38–51	26	25	9	35	7	23	7	16
52–72	26	25	10	38	11	35	20	45
HR,* beats/min								
Mean 93 ± 19								
Quartiles								
53–80	29	28	8	28	6	20	11	25
81–95	26	25	2	8	1	3	2	4
96–108	25	24	8	32	9	29	17	39
110–140	21	20	13	62	14	45	14	32
Unknown	3	3	1	33	1	3	0	0
SBP,* mmHg								
Mean 135 ± 22								
Quartiles								
70–125	26	25	13	50	12	39	18	41
126–138	27	26	4	15	3	10	3	7
139–150	28	27	6	21	9	29	14	32
151–180	21	20	8	38	6	19	9	20
Unknown	2	2	1	50	1	3	0	0
RR,* breaths/min								
Mean 17 ± 3								
Quartiles								
3–15	19	18	4	21	4	13	4	9
16	20	19	4	20	1	4	9	20
17–18	25	24	3	12	2	6	6	14
20–26	22	21	4	18	2	6	5	11
Unknown	18	17	17	94	22	71	20	46

*Denotes entry values taken from the run sheet.

Quartiles were established for age analysis. Race and age were not statistically different between those patients who received at least one LSI and those who received none, nor did race predispose to an LSI. Likewise, increasing patient age did not increase the frequency of an LSI in this sample/study. Of these 104 patients, 69% (72/104) did not receive an LSI. The other 31% (32/104) of patients received a total of 75 LSIs, of which 2% (2/104) of patients later died. Overall, 41% (31/75) of the LSIs were performed prehospital, and 59% (44/75) in the ED. Importantly, the demographics of the chosen population included HRs ranging from 53 to 140 beats/min, SBPs ranging from 70 to 180 mmHg, and various types of injuries and LSIs. This cohort provided the ECG morphology for HRV and HRC calculations.

Interventions performed on these 104 patients and classified as lifesaving by a multidisciplinary team of trauma experts are shown in Table 2. Interventions consisted of the following: 26 endotracheal intubations, 19 transfusions, 15 tube thoracostomies, seven cardiopulmonary resuscitations, one needle decompression, one pericardiocentesis, one cricothyrotomy, one thoracotomy, and four tourniquets.

Means of HRV and HRC statistics, SDs, and P values obtained via Wilcoxon tests for LSI and non-LSI patient groups are shown in Table 3. This table was used to confirm whether HRV and HRC can discriminate between the two patient groups; no other variables were included. Mean and minimum HRC (SampEn) values for patients who received at least one LSI were consistent with the fact that this group often has lower HRC than do patients who did not receive any LSI (14, 17). Mean HRV (HF-LF power spectrum ratio) values were not consistent with the fact that lower HRV is associated with increasing performance of LSIs.

For the first two sets of multivariate logistic regression analyses, results showed that increasing mean HR and decreasing total GCS score were associated with an increased risk for LSIs. Age, height, race, and weight were removed from the final models via backward elimination because they were not significantly associated with LSIs. In the model for vital signs alone (Table 4), odds ratios were 1.05 (95% confidence interval [CI], 1.03–1.09; $P < 0.0001$) for mean HR (per beats/min increase). In the model for vital signs and GCS scores (Table 5), odds

TABLE 2. **Lifesaving interventions**

LSIs	n	% (n/75)
Prehospital	31	41
Blood	3	4
Cardiopulmonary resuscitation	4	5
Chest tube	0	0
Intubation	22	29
Needle decompression	0	0
Pericardiocentesis	0	0
Surgical cricothyrotomy	0	0
Thoracotomy	0	0
Tourniquet	2	3
ED	44	59
Angio nonembolized	0	0
Angio embolized	0	0
Blood	16	22
Cardiopulmonary resuscitation	3	4
Cardioversion	0	0
Chest tube 1	10	14
Chest tube 2	5	7
Intubation	4	5
Needle decompression	1	1
Pericardiocentesis	1	1
Surgical cricothyrotomy	1	1
Thoracotomy	1	1
Tourniquet	2	3
Total	75	100

ratios were 1.05 (95% CI, 1.01–1.11; $P = 0.02$) for mean HR (per beats/min increase) and 0.68 (95% CI, 0.58–0.78; $P < 0.0001$) for total GCS score (per unit increase).

Inclusion of HRC in the multivariate logistic regression analyses showed that decreasing minimum HRC was also associated with an increased risk for LSIs. In the model for vital signs and HRC (Table 4), odds ratios were 1.05 (95% CI, 1.02–1.08; $P = 0.003$) for mean HR (per beats/min increase) and 0.00001 (95% CI, 0.00–0.05; $P = 0.012$) for minimum HRC (per unit increase). In the model for GCS scores and HRC (Table 5), odds ratios were 0.69 (95% CI, 0.59–0.78; $P < 0.0001$) for total GCS score (per unit increase) and 0.002 (95% CI, 0.00–11.29; $P = 0.16$) for minimum HRC (per unit increase).

Receiver operating characteristic curves (Fig. 1) demonstrated better identification for LSIs using HR and HRC (area under the curve [AUC] of 0.81) than using HR alone (AUC of 0.73). Likewise, ROC curves (Fig. 2) demonstrated better identification for LSIs using total GCS score and HRC (AUC of 0.94) than using total GCS score and HR (AUC of 0.92).

Importantly, multiple ML models were developed, trained, and compared for the outcomes of at least one LSI and no LSIs. A model consisting of a multilayer perceptron with three inputs (mean HR, total GCS score, minimum HRC) and three hidden nodes yielded the best results (Fig. 3). Receiver operating characteristic curves (Fig. 4) demonstrated that an ML model using HR, total GCS score, and HRC (AUC of 0.99) had superior performance over multivariate logistic regression models (Figs. 1 and 2) for identifying the needs for LSIs in trauma patients.

TABLE 3. **Comparison of HRV and HRC values between patient groups**

	LSI, no. patients (n = 32)		NLSI, no. patients (n = 72)		
Measure	Mean	SD	Mean	SD	P
Mean HRV	1.527	0.526	1.527	0.366	0.701
Maximum HRV	5.061	1.477	4.091	1.168	0.001
Minimum HRV	0.452	0.217	0.423	0.182	0.660
Mean HRC	0.595	0.223	0.675	0.210	0.158
Maximum HRC	1.230	0.222	1.218	0.169	0.206
Minimum HRC	0.045	0.068	0.111	0.100	0.001

NLSI indicates nonlifesaving intervention.

TABLE 4. **Logistic regression models with various risk factors (excluding GCS) for LSIs**

Variable	Odds ratio for LSIs (95% CI)*	P
Mean HR	1.05 (1.03–1.09)	<0.0001
With HRC		
Mean HR	1.05 (1.02–1.08)	0.003
Minimum HRC	0.00001 (0.00–0.05)	0.012

*Odds ratios for measurements reflect per-unit increase.

DISCUSSION

This study examined the utility of standard vital signs, HRV, HRC, and ML for predicting the need for LSIs in trauma patients by comparing the performance of multivariate logistic regression identification versus ML technologies. Previous studies analyzed only traditional vital signs (29) or a combination of HRV, HRC, and ML (17) for discriminating between LSI and non-LSI patients. In the former case, neither HRC nor ML was used for identifying LSI patients, and in the latter case, neither standard vital signs nor GCS scores were used for identifying LSI patients, resulting in models achieving ROC AUC of no more than 0.868. Likewise, no comparisons were performed using different models. Baxt and colleagues used the motor component of the GCS score for analysis of trauma patients, but in the context of triage, not the identification of trauma patients receiving LSIs (30). Recent work reported the

TABLE 5. **Logistic regression models with various risk factors (including GCS) for LSIs**

Variable	Odds ratio for LSIs (95% CI)*	P
Mean HR	1.05 (1.01–1.11)	0.02
Total GCS score	0.68 (0.58–0.78)	<0.0001
With HRC		
Total GCS score	0.69 (0.59–0.78)	<0.0001
Minimum HRC	0.002 (0.00–11.29)	0.16

*Odds ratios for measurements reflect per-unit increase.

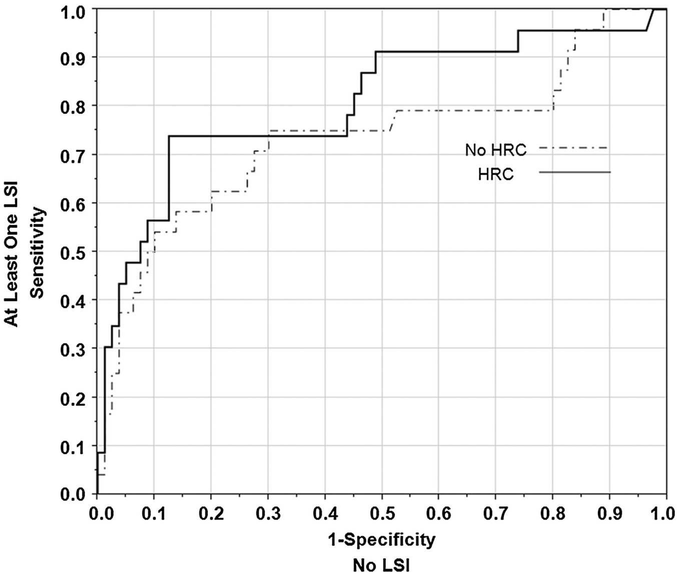

Fig. 1. **Receiver operating characteristic curves for models (excluding GCS scores)**. Receiver operating characteristic curves were obtained to examine the discriminating power of multivariate logistic regression models (excluding GCS scores) for the outcome of at least one LSI in 104 subjects. The curves demonstrated better LSI identification for models using both vital signs and HRC (AUC of 0.81) than for models using only vital signs (AUC of 0.73).

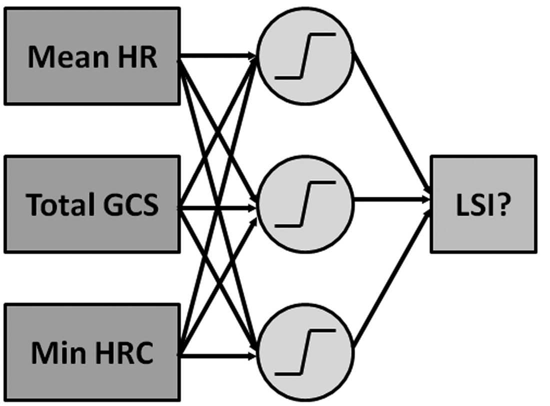

Fig. 3. **Three-layer perceptron model**. After developing and training multiple ML models for the outcomes of at least one LSI and no LSIs, one best model was selected. The model consisted of a multilayer perceptron with three inputs (mean HR, GCS score, and minimum HRC [min HRC]) and three hidden nodes for identifying LSI needs.

development and validation of a real-time LSI prediction system but excluded HRV and HRC analyses (6). A novelty of this study was the exploration of traditional and new vital signs for predicting LSIs, showing that an ML model was superior in performance over multivariate logistic regression models.

Based on our results, statistics derived from the WVSM data confirmed that HRC alone may be able to discriminate between those patients who received one or more LSIs and those who received none. However, because of noise in the ECG waveforms and the sensitivity of HRV to noise, this study could not show that HRV differs between LSI and non-LSI patients. In this case, noise in the ECG waveforms prevented accurate calculation of HRV (HF-LF power spectrum ratio) in many patients.

In this study, increasing HR mean increased the odds of an LSI by approximately 5%. In addition, the presence of decreasing minimum HRC without GCS scores in the multivariate logistic regression model increased the odds of an LSI by more than 1,000%. These findings appeared to be similar to previous work, which reported that lower HRC in patients could lead to more expeditious identification of battlefield casualties in need of LSIs (14, 17). When GCS scores were incorporated into the logistic regression models, the presence of HRC was not as significant. Again, these findings agreed with earlier work, which concluded that GCS scores reliably identify the need for prehospital LSIs in trauma patients (29). These results

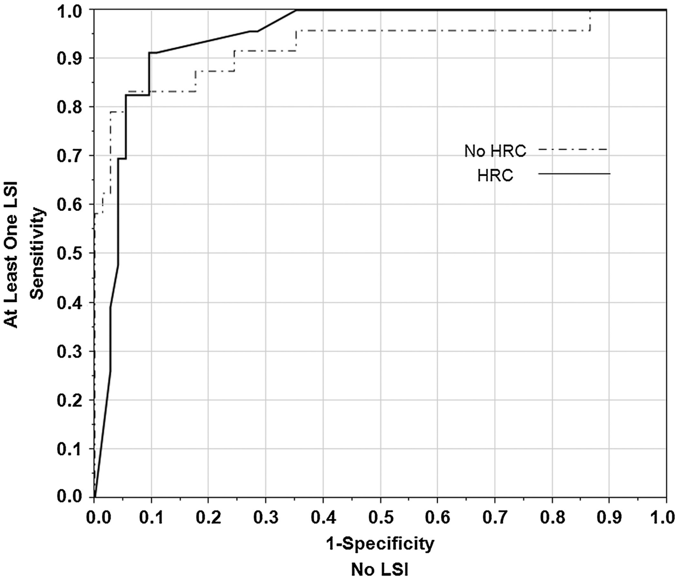

Fig. 2. **Receiver operating characteristic curves for models (including GCS scores)**. Receiver operating characteristic curves were obtained to examine the discriminating power of multivariate logistic regression models (including GCS scores) for the outcome of at least one LSI in 104 subjects. The curves demonstrated better LSI identification for models using vital signs, GCS scores, and HRC (AUC of 0.94) than for models using only vital signs and GCS scores (AUC of 0.92).

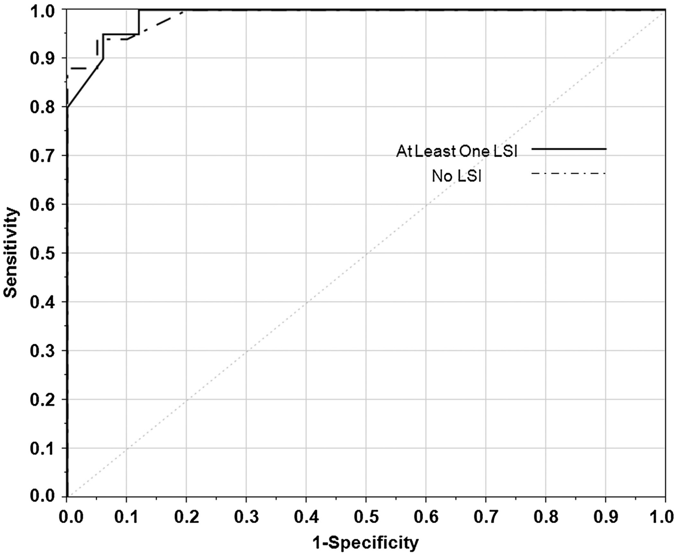

Fig. 4. **Receiver operating characteristic curves for a three-layer perceptron model**. Receiver operating characteristic curves were obtained to examine the discriminating power of an ML model (three-layer perceptron with three inputs [mean HR, GCS score, and HRC] and three hidden nodes) for the outcome of at least one LSI and no LSIs in 104 subjects. The curves demonstrated superior LSI identification performance (AUC of 0.99) for at least one LSI and no LSIs.

also appeared similar to the work of Baxt and colleagues (30), who used the GCS motor score to develop a triage rule.

Importantly, this study demonstrated that multivariate logistic regression models incorporating HRC could increase the LSI identification accuracy for this cohort. The hypothesis that an ML model utilizing a combination of vital signs, HRV, and HRC to identify LSI needs could outperform multivariate logistic regression models utilizing a similar combination was shown through a comparison of ROC curves and AUC results.

It is important to point out why models utilizing and not utilizing total GCS scores were considered separately in this study. GCS scores when available are convenient but do not always support the concepts of automation and continuous data analysis, especially within a battlefield environment. In other words, GCS scores require physical examination of the patient and are not always available when basing treatment decisions based solely on electronic data (e.g., evacuation). Automation and continuous data analysis have many potential implications for both military and civilian trauma care. Constant physiologic observations and data could enhance the medic's ability to assess and treat battlefield and civilian injuries. In addition, continuous physiologic data could improve triage and treatment of trauma patients for both military and civilian trauma centers (2, 5, 6). This study showed that models not utilizing GCS scores were still able to perform LSI identifications with greater than 80% accuracy. By integrating surrogate injury scores (suitable for continuous data analysis) along with vital signs and HRC into this study's models, it is possible to preclude use of GCS scores while increasing LSI identification accuracy. This hypothesis could be a future study using either the same WVSM data set or a larger data set reflecting blunt and penetrating injuries.

The results of this study suggest that HRV for a trauma patient cohort may require more careful examination of underlying waveforms before use in a clinical setting. By screening out unreliable ECG waveforms and resulting HRV measurements from further analysis (17), this study might have confirmed that HRV can discriminate between those patients who received one or more LSIs and those who received none. This supports evidence that HRV is lower in LSI patients than in non-LSI patients (17).

A major implication of this study was that development of CDS systems should utilize vital signs, HRC, ML, and other information in order to achieve more accurate diagnostic capabilities. In addition, HRC may be more suitable for clinical use when analyzed in conjunction with vital signs. Future studies may include indicators of numeric and waveform data quality to provide a more comprehensive model for predictions of outcomes in trauma patients.

Limitations

This study had several limitations. The size of the data set was small; i.e., it contained data from 104 patients in total. Moreover, the results were preliminary because of the data set size and the criteria for selecting the data. No injury severity scores were recorded. Lifesaving interventions were recorded only when the nurse/paramedic manually pressed a button on the WVSM data-capture-and-display interface. Because of this limitation, the study suffered from scarcity of recorded times of LSIs needed to validate model development and performance. Lastly, this study did not consider separate analyses for examining the discriminating power of the models for the outcome of at least one prehospital LSI or one ED LSI, nor did models incorporate trends to determine their utility. A strategy similar to this study could be applied to perform these analyses in the future.

In summary, this study showed the power of vital sign measurements, HRC, and ML to identify whether LSIs were performed in 104 trauma patients with blunt or penetrating injuries. An ML model was shown to be superior over various logistic regression models. Development of CDS systems should utilize vital signs, HRC, and ML in order to achieve more accurate diagnostic capabilities, such as identification of needs for LSIs in trauma patients.

ACKNOWLEDGMENTS

The authors acknowledge the expertise, dedication, and professionalism of the emergency medical services paramedics, nurses, and staff in Houston who performed the patient care and Denise Hinds, Timothy Welch, and Jeannette Podbielski (the University of Texas Health Science Center at Houston, Texas).

REFERENCES

1. Yilmaz T, Foster R, Hao Y: Detecting vital signs with wearable wireless sensors. *Sensors* 10:10837–10862, 2010.
2. Salinas J, Nguyen R, Darrah MI, Kramer GA, Serio-Melvin ML, Mann EA, Wolf SE, Chung KK, Renz EM, Cancio LC: Advanced monitoring and decision support for battlefield critical care environment. *US Army Med Dep J* Apr-Jun:73–81, 2011.
3. Garg AX, Adhikari NKJ, McDonald H, Rosas-Arellano MP, Devereaux PJ, Beyene J, Sam J, Haynes RB: Effects of computerized clinical decision support systems on practitioner performance and patient outcomes. A systematic review. *JAMA* 293:1223–1238, 2005.
4. Kawamoto K, Houlihan CA, Balas EA, Lobach DF: Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. *BMJ* 10:1–8, 2005.
5. Shoemaker WC, Wo CC, Lu K, Chien LC, Rhee P, Bayard D, Demetriades D, Jelliffe RW: Noninvasive hemodynamic monitoring for combat casualties. *Mil Med* 171:813–820, 2006.
6. Liu NT, Holcomb JB, Wade CE, Batchinsky AI, Cancio LC, Darrah MI, Salinas J: Development and validation of a machine learning algorithm and hybrid system to predict the need for life-saving interventions in trauma patients. *Med Biol Comput Eng* 52:193–203, 2014.
7. Chen L, Reisner AT, Gribok A, Reifman J: Exploration of prehospital vital sign trends for the identification of trauma outcomes. *Prehosp Emerg Care* 13: 286–294, 2009.
8. Pickering TG, Shimbo D, Hass D: Ambulatory blood-pressure monitoring. *N Engl J Med* 354:2368–2374, 2006.
9. Low RB, Martin D: Accuracy of blood pressure measurements made aboard helicopters. *Ann Emerg Med* 17:604–612, 1988.
10. Garner DC: Noise in medical helicopters. *JAMA* 266:515–516, 1991.
11. Jones DW, Appel LJ, Sheps SG, Roccella EJ, Lenfant C: Measuring blood pressure accurately: new and persistent challenges. *JAMA* 289:1027–1030, 2003.
12. Lovett PB, Buchwald JM, Sturmann K, Bijur P: The vexatious vital: neither clinical measurements by nurses nor an electronic monitor provides accurate measurements of respiratory rate in triage. *Ann Emerg Med* 45:68–76, 2005.
13. Batchinsky AI, Cooke WH, Kuusela T, Cancio LC: Loss of complexity characterizes the heart-rate response to experimental hemorrhagic shock in swine. *Crit Care Med* 35:519–525, 2007.
14. Cancio LC, Batchinsky AI, Salinas J, Kuusela T, Convertino VA, Wade CE, Holcomb JB: Heart-rate complexity for identification of prehospital lifesaving interventions in trauma patients. *J Trauma* 65:813–819, 2008.
15. Norris PR, Anderson SM, Jenkins JM, Williams AE, Morris JA Jr: Heart rate multiscale entropy at three hours predicts hospital mortality in 3,154 trauma patients. *Shock* 30:17–22, 2008.
16. Batchinswky AL, Salinas J, Kuusela T, Necsoiu C, Jones J, Cancio LC: Rapid prediction of trauma patient survival by analysis of heart rate complexity: impact of reducing data set size. *Shock* 32:565–571, 2009.

17. Batchinsky AI, Salinas J, Jones JA, Necsoiu C, Cancio LC: Identifying the need to perform life-saving interventions in trauma patients using new vital signs and artificial neural networks. *Lect Notes Comput Sc* 5651:390–394, 2009.
18. Batchinsky AI, Skinner J, Necsoiu C, Jordan BS, Weiss D, Cancio LC: New measures of heart-rate complexity: effect of chest trauma and hemorrhage. *J Trauma* 68:1178–1185, 2010.
19. Clemens MG: The data sets needed for analysis of heart-rate complexity to identify trauma patients with potentially lethal injuries. *Shock* 33:1–2, 2010.
20. Liu NT, Cancio LC, Salinas J, Batchinsky AI: Reliable real-time calculation of heart-rate complexity in critically ill patients using multiple noisy waveform sources. *J Clin Monit Comput* 28:123–131, 2014.
21. Ryan KL, Rickards CA, Ludwig DA, Convertino VA: Tracking central hypovolemia with ECG in humans: cautions for the use of heart period variability in patient monitoring. *Shock* 33:583–589, 2010.
22. Liu NT, Batchinsky AI, Cancio LC, Baker WL, Salinas J: Development and validation of a novel fusion algorithm for continuous, accurate, and automated R-wave detection and calculation of signal-derived metrics. *J Crit Care* 28:885.e9–885.e18, 2013.
23. Liu NT, Batchinsky AI, Cancio LC, Salinas J: The impact of noise on the reliability of heart-rate variability and complexity analysis in trauma patients. *Comput Biol Med* 43:1955–1964, 2013.
24. Heldt T, Long B, Verghese GC, Szolovits P, Mark RG: Integrating data, models, and reasoning in critical care. *Proc 28th IEEE EMBS Annu Int Conf* 1:350–353, 2006.
25. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Heart rate variability. Standards of measurement, physiological interpretation and clinical use. *Circulation* 93: 1043–1065, 1996.
26. Ellenby MS, McNames J, Lai S, McDonald BA, Krieger D, Sclabassi RJ, Goldstin B: Uncoupling and recoupling of autonomic regulation of the heart beat in pediatric septic shock. *Shock* 16:274–277, 2001.
27. Richman JS, Moorman JR: Physiological time series analysis using approximate entropy and sample entropy. *Am J Physiol Heart Circ Physiol* 278: H2039–H2049, 2000.
28. Lake DE, Richman JS, Griffin MP, Moorman JR: Sample entropy analysis of neonatal heart rate variability. *Am J Physiol Regul Integr Comp Physiol* 283: R789–R797, 2002.
29. Holcomb JB, Salinas J, McManus JM, Miller CC, Cooke WH, Convertino VA: Manual vital signs reliably identify need for life-saving interventions in trauma patients. *J Trauma* 59:821–829, 2005.
30. Baxt WG, Jones G, Fortlage D: The trauma triage rule: a new, resource-based approach to the prehospital identification of major trauma victims. *Ann Emerg Med* 19:1401–1406, 1990.